



# (12) 发明专利

(10) 授权公告号 CN 101388988 B

(45) 授权公告日 2012. 05. 23

(21) 申请号 200810173720. 2

(51) Int. Cl.

(22) 申请日 2008. 06. 10

H04N 5/76 (2006. 01)

H04N 13/00 (2006. 01)

(30) 优先权数据

56343/07 2007. 06. 08 KR

59662/07 2007. 06. 18 KR

审查员 张春

(73) 专利权人 三星电子株式会社

地址 韩国京畿道

专利权人 庆熙大学校产学协力团

(72) 发明人 郑度泳 朴泰诚 吴润济 徐德荣

朴光勋 金奎宪 李润辰

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 邵亚丽

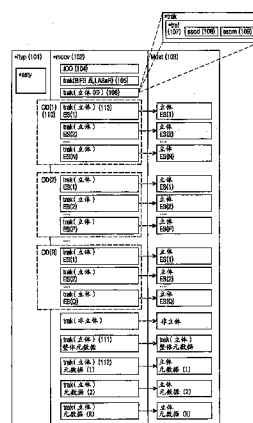
权利要求书 2 页 说明书 10 页 附图 11 页

## (54) 发明名称

用于记录三维视频数据的方法

## (57) 摘要

公开了一种用于将包括已经通过转换由支持三维图像拍摄的设备输入的图像而生成的三维图像的场景存储到记录介质中的方法,包括:在预定的文件类型字段中,装载指示单一立体图像以及合成图像之一的图像类型信息;在配置性信息容器字段中,装载指示该三维图像中包括的多个媒体对象之间的时间和空间关系的场景描述符信息、指示每个媒体对象的对象属性和所述每个媒体对象的编码流之间的配置性关系的对象配置性信息、编码流的配置性信息、多个媒体对象之间的第一元数据、和每个媒体对象与其编码流之间的第二元数据;以及在媒体数据容器字段中,装载每个媒体对象的比特流、用于第一元数据的比特流、和用于第二元数据的比特流。



1. 一种用于将已经通过转换由支持三维图像数据拍摄的设备输入的图像数据而生成的三维图像存储到记录介质中的方法,该方法包括:

将通过多个镜头输入的颜色信号转换成电信号,并且从转换的电信号中生成左图像数据和右图像数据;

对左图像数据和右图像数据进行编码;

生成包括左图像数据和右图像数据的三维图像数据;以及

将该三维图像数据存储在存储器中;

其中存储该三维图像数据的步骤包括:

在预定的文件类型字段中,装载指示单一立体图像、以及包括单视场图像和立体图像的合成图像之一的图像类型信息;

在配置性信息容器字段中,装载指示该三维图像中包括的多个媒体对象之间的时间和空间关系的场景描述符信息、指示所述多个媒体对象的每个媒体对象的对象属性和多个媒体对象的编码流之间的配置性关系的对象配置性信息、编码流的配置性信息、媒体对象之间的第一元数据、和每个媒体对象与其编码流之间的第二元数据;以及

在媒体数据容器字段中,装载该三维图像中包括的每个媒体对象的比特流、用于第一元数据的比特流、和用于第二元数据的比特流。

2. 根据权利要求1的方法,其中所述对象配置性信息包括:

指示场景中包括的对象的数目的信息;

指示对象是否是立体图像的信息;

指示对象是否需要元数据的信息;

指示对象中包括的编码流的数目的信息;

当元数据被应用于对象时,指示包括要被应用的元数据的轨道的信息;以及

指示包括编码流的轨道的信息。

3. 根据权利要求2的方法,其中所述对象配置性信息进一步包括:指示包括可应用于整个立体图像的元数据的轨道的信息。

4. 根据权利要求1的方法,存储该三维图像数据的步骤进一步包括:装载合成图像的配置性信息。

5. 一种用于将已经通过转换由支持三维图像数据拍摄的设备输入的图像数据而生成的三维图像存储到记录介质中的方法,该方法包括:

将通过多个镜头输入的颜色信号转换成电信号,并且从转换的电信号中生成左图像数据和右图像数据;

对左图像数据和右图像数据进行编码;

生成包括左图像数据和右图像数据的三维图像数据;以及

将该三维图像数据存储在存储器中;

其中存储该三维图像数据的步骤包括:

在预定的文件类型字段中,装载指示单一立体图像、以及包括单视场图像和立体图像的合成图像之一的图像类型信息;

在配置性信息容器字段中,装载指示多个媒体对象的每个媒体对象的对象属性和该三维图像数据中包括的每个媒体对象的编码流之间的配置性关系的对象配置性信息、每个编

码流的配置性信息、媒体对象之间的第一元数据、和每个媒体对象与其编码流之间的第二元数据；以及

在媒体数据容器字段中，装载该三维图像中包括的每个媒体对象的比特流、用于第一元数据的比特流、和用于第二元数据的比特流，

其中，可选择性地在配置性信息容器字段中，装载指示该三维图像中包括的所述多个媒体对象的媒体对象之间的时间和空间关系的场景的场景描述符信息，以及  
编码流的配置性信息包括关于是否装载场景描述符信息的信息。

6. 根据权利要求 5 的方法，其中所述场景描述符信息包括：

初始场景描述符信息字段，包含关于初始访问指针和场景描述的信息；

场景信息字段，包含关于该场景中包括的多个媒体对象的配置性关系的信息；以及

对象配置性信息字段，包含关于该场景中包括的多个媒体对象的每个媒体对象的属性的信息。

7. 根据权利要求 6 的方法，其中，当该配置性信息容器字段包括初始场景描述符信息时，编码流的配置性信息包含关于该初始场景描述符信息中包括的多个媒体对象的每个媒体对象的解码器信息。

8. 根据权利要求 6 的方法，其中，当该配置性信息容器字段不包括初始场景描述符信息时，编码流的配置性信息包含关于该立体图像中包括的多个媒体对象的每个媒体对象的解码器信息。

9. 根据权利要求 8 的方法，其中关于该立体图像中包括的多个媒体对象的每个媒体对象的解码器信息包括：编码流的类型、编码流的媒体类型、简档信息、和关于图像的详细信息。

10. 根据权利要求 9 的方法，其中该媒体类型与基于由 ISO/IEC 14496MPEG-4 系统定义的对象类型指示 (OTI) 设置的信息对应。

11. 根据权利要求 5 的方法，其中，该第一元数据包括关于编码流的比特流的长度和该编码流的标识符的信息，而该第二元数据包括关于每个编码流的名称和内容类型的信息、及编码信息。

## 用于记录三维视频数据的方法

### 技术领域

[0001] 本发明涉及一种用于记录三维视频数据的记录介质和方法,且更具体地,涉及一种用于以具有与 MPEG14496-12ISO 基本媒体格式兼容的结构的格式来生成和存储三维视频数据的方法。

### 背景技术

[0002] 典型的媒体文件格式基本地包括:用于记录媒体信息的报头部分和用于存储压缩后的媒体数据的视频数据部分。然而,尽管典型的媒体文件格式适于存储简单的视频数据,但其不适于作为用于存储各种类型的媒体的一般结构。

[0003] 在这点上,作为国际标准化组织的运动图像专家组(MPEG)在标准 ISO/IEC14496 中定义 MPEG-4ISO 基本媒体文件格式作为可普遍应用于各种应用的基本文件格式。已经设计 ISO 基本媒体文件格式来在多个容器中分层次地存储压缩后的媒体流和关于该媒体流的数据(例如,诸如配置性信息)。虽然 ISO 基本媒体文件格式并没有基本地定义编码或解码方案,但却定义了用于有效存储编码或解码媒体流的基本结构。

[0004]

| 部分   | 编号              | 标题 | 描述                                                                                              |
|------|-----------------|----|-------------------------------------------------------------------------------------------------|
| 部分 1 | ISO/IEC 14496-1 | 系统 | 描述视频和音频的同步和多路复用。例如,传输流。                                                                         |
| 部分 2 | ISO/IEC 14496-2 | 可视 | 用于可视数据(视频、静止纹理、合成图像等)的压缩编解码器。部分 2 中的多个“简档(profiles)”中的一个 是高级简单简档(advanced simple profile, ASP)。 |
| 部分 3 | ISO/IEC 14496-3 | 音频 | 用于对音频信号进行感知编码的压缩编解码器的集合,包括高级音频编码(AAC)和其它音频/语音编码工具的                                              |

[0005]

|       |                  |                    |                                    |
|-------|------------------|--------------------|------------------------------------|
|       |                  |                    | 一些变化。                              |
| 部分 4  | ISO/IEC 14496-4  | 一致性                | 描述用于测试与标准的其它部分的一致性的过程。             |
| 部分 5  | ISO/IEC 14496-5  | 参考软件               | 提供用于展示和阐明标准的其它部分的软件。               |
| 部分 6  | ISO/IEC 14496-6  | 传递多媒体集成框架(DMIF)。   |                                    |
| 部分 7  | ISO/IEC 14496-7  | 优化参考软件             | 提供如何进行改善的实现的示例(例如,关于 Part5)。       |
| 部分 8  | ISO/IEC 14496-8  | IP 网络上的承载          | 指定用于在 IP 网络上承载 MPEG-4 内容的方法。       |
| 部分 9  | ISO/IEC 14496-9  | 参考硬件               | 提供用于展示如何实现标准其它部分的硬件设计。             |
| 部分 10 | ISO/IEC 14496-10 | 高级视频编码(AVC)        | 技术上等同于 ITU-T H.264 标准的用于视频信号的编解码器。 |
| 部分 11 | ISO/IEC 14496-11 | 场景描述和应用引擎(“BIFS”)  | 可被应用于具有多个简档的丰富的交互内容,包括 2D 和 3D 版本。 |
| 部分 12 | ISO/IEC 14496-12 | ISO 基本媒体文件格式       | 用于存储媒体内容的文件格式。                     |
| 部分 13 | ISO/IEC 14496-13 | 知识产权管理和保护(IPMP)扩展。 |                                    |
| 部分 14 | ISO/IEC 14496-14 | MPEG-4 文件格式        | 基于部分 12 的用于 MPEG-4 内容的指定的容器文件格式。   |
| 部分 15 | ISO/IEC 14496-15 | AVC 文件格式           | 用于基于部分 12 来存储部分 10 视频。             |
| 部分 16 | ISO/IEC 14496-16 | 动画框架扩展(AFX)。       |                                    |
| 部分 17 | ISO/IEC 14496-17 | 定时文本字幕格式。          |                                    |

[0006]

|       |                        |                            |                                                    |
|-------|------------------------|----------------------------|----------------------------------------------------|
| 部分 18 | ISO/IEC 14496-18       | 字体压缩和流（用于 OpenType 字体）。    |                                                    |
| 部分 19 | I S O / I E C 14496-19 | 合成纹理流。                     |                                                    |
| 部分 20 | ISO/IEC 14496-20       | 轻量级应用场景表示 (LASER)。         |                                                    |
| 部分 21 | ISO/IEC 14496-21       | MPEG-J 图形框架扩展 (GFX)        | ( 还未完成 - 在 2005 年 7 月处于 “FCD” 阶段,2006 年 1 月 FDIS)。 |
| 部分 22 | I S O / I E C 14496-22 | 基于 OpenType 的开放字体格式 (OFFS) | ( 还未完成 - 在 2005 年 7 月达到 “CD” 阶段 )                  |
| 部分 23 | ISO/IEC 14496-23       | 符号音乐表示 (SMR)               | ( 还未完成 - 在 2006 年 10 月达到 “FCD” 阶段 )                |

[0007] 然而，具有为了存储二维视频所定义的结构的那种媒体文件格式不适于存储包括立体图像和各种媒体数据的三维视频。因而，需要考虑立体图像的结构而定义的数据存储格式。同时地，需要这种立体图像具有与作为国际标准的 ISO 基本媒体文件格式兼容的数据存储格式，以便其可以应用于各种多媒体应用。具体地，数据存储格式必须与图 7A 中所参考的 14496-12ISO 基本媒体文件格式兼容。

发明内容

[0008] 相应地，已经作出了本发明以解决现有技术中出现的上述问题，并且本发明提供了附加优点。特别地，本发明提供一种包括关于立体图像的配置的信息且与作为国际标准的 ISO 基本媒体文件格式兼容的数据存储格式、和一种用于生成该数据存储格式的方法。

[0009] 根据本发明的一方面，提供了一种用于将已经通过转换由支持三维图像拍摄的设备输入的图像而生成的三维图像存储到记录介质中的方法，该方法包括步骤：在预定的文件类型字段中，装载指示单一立体图像、以及包括单视场图像和立体图像的合成图像之一的图像类型信息；在配置性信息容器字段中，装载指示该三维图像中包括的多个媒体对象之间的时间和空间关系的场景描述符信息、指示所述多个媒体对象的每个媒体对象的对象属性和多个媒体对象的多个编码流之间的配置性关系的对象配置性信息、多个编码流的每个流的配置性信息、多个媒体对象之间的第一元数据、及所述多个媒体对象的每个媒体对象和所述多个编码流的一个编码流之间的第二元数据；以及在媒体数据容器字段中，装载该三维图像中包括的多个媒体对象的每个媒体对象的比特流、用于第一元数据的比特流、和用于第二元数据的比特流。

[0010] 所述对象配置性信息包括：指示场景中包括的多个媒体对象的对象数目的信息；指示所述多个对象的每个对象是否是立体图像的信息；指示所述多个对象中的每个对象是否需要元数据的信息；指示每个对象中包括的多个编码流的数目的信息；当元数据被应用于所述多个对象的对象时，指示包括要被应用的元数据的轨道的信息；以及指示包括所述多个编码流的编码流的轨道的信息。

[0011] 所述对象配置性信息可进一步包括：指示包括可应用于整个立体图像的元数据的轨道的信息。

[0012] 根据本发明的另一方面，提供了一种其中装载三维图像的计算机可读介质，该介质包括：文件类型字段，包含指示单一立体图像、以及包括单视场图像和立体图像的合成图像之一的图像类型信息；配置性信息容器字段，包含指示该三维图像中包括的多个媒体对象之间的时间和空间关系的场景描述符信息、指示所述多个媒体对象的每个媒体对象的对象属性和多个媒体对象的多个编码流之间的配置性关系的对象配置性信息、多个编码流的配置性信息、多个媒体对象之间的第一元数据、及多个媒体对象的每个媒体对象和多个编码流的对应的编码流之间的第二元数据；以及媒体数据容器字段，包含该三维图像中包括

的每个媒体对象的比特流、用于第一元数据的比特流、和用于第二元数据的比特流。

[0013] 根据本发明的另一方面,提供了一种用于将已经通过转换由支持三维图像拍摄的设备输入的图像而生成的三维图像存储到记录介质中的方法,该方法包括:在预定的文件类型字段中,装载指示单一立体图像、以及包括单视场图像和立体图像的合成图像之一的图像类型信息;在配置性信息容器字段中,装载指示多个媒体对象的每个媒体对象的对象属性和包括与该三维图像中包括的多个媒体对象的每个媒体对象对应的编码流的多个编码流的配置性关系的对象配置性信息、多个编码流的每个编码流的配置性信息、所述多个媒体对象的媒体对象之间的第一元数据、和每个媒体对象与其对应的编码流之间的第二元数据;以及在媒体数据容器字段中,装载该三维图像中包括的每个媒体对象的比特流、用于第一元数据的比特流、和用于第二元数据的比特流,其中,可选择性地在配置性信息容器字段中,装载指示该三维图像中包括的所述多个媒体对象的媒体对象之间的时间和空间关系的场景描述符信息,以及多个编码流的配置性信息包括关于是否装载场景描述符信息的信息。

[0014] 场景的所述场景描述符信息包括:初始场景描述符信息字段,包含关于初始访问指针和场景描述的信息;场景信息字段,包含关于该场景中包括的多个媒体对象的配置性关系的信息;以及对象配置性信息字段,包含关于所述多个媒体对象的每个媒体对象的属性的信息。

[0015] 当该配置性信息容器字段包括初始场景描述符信息时,编码流的配置性信息包含关于该初始场景描述符信息中包括的多个媒体对象的解码器信息。

[0016] 当该配置性信息容器字段不包括初始场景描述符信息时,编码流的配置性信息包含关于该立体图像中包括的多个媒体对象的解码器信息。

[0017] 该媒体类型可以是基于 MPEG-4 系统中定义的对象类型标志 (OTI) 设置的信息。

[0018] 该第一元数据包括关于编码流的比特流的长度和该编码流的标识符的信息,而该第二元数据包括关于每个编码流的名称和内容类型的信息,及编码信息。

[0019] 根据本发明的另一方面,提供了一种其中装载三维图像的计算机可读介质,该介质包括:预定的文件类型字段,包含指示单一立体图像、以及包括单视场图像和立体图像的合成图像之一的图像类型信息;配置性信息容器字段,包含指示多个媒体对象的每个媒体对象的对象属性和该三维图像中包括的媒体对象的多个编码流之间的配置性关系的对象配置性信息、编码流的配置性信息、媒体对象之间的第一元数据、和每个媒体对象及编码流之间的第二元数据;以及媒体数据容器字段,包含该三维图像中包括的每个媒体对象的比特流、用于第一元数据的比特流、和用于第二元数据的比特流,其中,可选择性地在配置性信息容器字段中,装载指示该三维图像中包括的多个媒体对象的媒体对象之间的时间和空间关系的场景描述符信息,以及编码流的配置性信息包括关于是否装载场景描述符信息的信息。

## 附图说明

[0020] 通过接下来结合附图的详细描述,本发明的上述和其它方面、特征和优点将更加明显,其中:

[0021] 图 1 是应用了本发明的立体图像拍摄设备的方框图;

[0022] 图 2 图示了根据本发明实施例的三维图像中包括的立体图像的示例;

- [0023] 图 3 图示了根据本发明实施例的三维图像数据的存储格式；
- [0024] 图 4 图示了根据本发明实施例的用于生成存储格式中包括的 ssod 字段的程序的示例；
- [0025] 图 5 图示了根据本发明实施例的用于生成存储格式中包括的 ssod 字段的程序的示例；
- [0026] 图 6 图示了根据本发明另一实施例的用于存储三维图像数据的格式；
- [0027] 图 7A、7B、7C、和 7D 图示了根据本发明另一实施例的用于生成存储格式中包括的 mpss 字段的程序的示例；
- [0028] 图 8 图示了根据本发明另一实施例的在包括场景描述符信息的情况下的数据存储格式的结构；以及
- [0029] 图 9 图示了根据本发明另一实施例的在不包括场景描述符信息的情况下的数据存储格式的结构。

### 具体实施方式

[0030] 在下文中,将参考附图描述本发明的实施例。为了清楚和简明的目的,将省略对在这里所合并的已知功能和配置的详细描述,因为它可造成本发明的主题有些不清楚。

[0031] 本发明最新提供了作为在仅一个 MS 中使用的特定 CID 的专用 (private) 传输 CID。在参考图 2 描述用于分配传输 CID 的传统方案之后,将参考图 3 更详细地描述该专用传输 CID。

[0032] 图 1 是应用了本发明的立体图像拍摄设备的方框图。

[0033] 立体图像拍摄设备包括:图像输入单元 10、编码器 20、三维图像发生器 30、数据存储控制器 40、存储器 50、和显示器 60。

[0034] 图像输入单元 10 与用于输入生成立体图像必需的多个图像的装置对应,且包括:通过其以各种角度来投影从对象反射的特定波长的颜色信号的多个镜头、以及用于将通过镜头输入的颜色信号转化为电信号的多个传感器模块(包括传感器的模块,例如:CCD, CMOS 等)。此外,图像输入单元 10 还包括视频信号处理设备,其将从多个传感器模块输入的数据转换为包括时间/空间信息的图像信号,从而生成和传感器模块一样多条的图像数据。例如,当立体图像包括左侧图像和右侧图像的媒体数据的组合时,图像输入单元 10 包括用于拍摄对象左侧部分的镜头和传感器模块、以及用于拍摄对象右侧部分的镜头和传感器模块,这样,图像输入单元 10 通过使用从所述两个传感器模块输入的数据,来生成包括左侧图像和右侧图像的三维图像。

[0035] 编码器 20 编码从图像输入单元 10 输出的图像数据(包括左侧图像数据和右侧图像数据),并且可使用典型的动态图像编码方案(例如,MPEG-1、MPEG-2、MPEG-3、MPEG-4、H.264 方案)。

[0036] 三维图像发生器 30 考虑与拍摄对象的距离、拍摄角度、拍摄设备的运动等,从多条编码后的媒体数据生成三维图像数据。

[0037] 其间,如果用户长时间观看仅包括立体图像的三维图像,则用户的眼睛比当用户观看单视场(monoscopic)图像时感觉更疲劳。由于这个原因,用户很难长时间观看仅包括立体图像的三维图像。因而,为了减少眼睛的疲劳并确保长时间观看三维图像,优选地,三维图像数据包括组合单视场图像和立体图像的合成图像,而不是仅包括立体图像。可以在拍摄设备的用户的控制下或者根据在拍摄设备中预先设定的预定条件下,恰当地组合立体图像和单视场图像。所述预定条件可被设计来以预定的时间间隔周期性地拍摄立体图像或单视场图像,或者根据所拍摄对象的特性或类型来可选择性地拍摄立体图像和单视场图像。

[0038] 而且,立体图像可进一步包括:诸如话音信息和文本信息的各种媒体对象,并且立体图像拍摄设备可包括:用于将所生成的话音信号转换为电子信号的传声器、用于压缩音频信号的话音压缩单元、用于从外部输入话音信息的与话音输入终端连接的话音信息输入接口。此外,立体图像拍摄设备还可包括用于从外部输入文本信息的文本信息输入接口。

[0039] 此外,三维图像发生器 30 还向数据存储控制器 40 提供包括三维图像的图像、话音、和文本信息的三维图像配置性信息。

[0040] 数据存储控制器 40 在存储器 50 中存储从所述三维图像配置性信息所生成的三维图像数据以及从三维图像发生器 30 所提供的三维图像数据。此外,显示器 60 输出由三维图像发生器 30 生成的三维图像。

[0041] 此外,根据本发明的立体图像拍摄设备可包括通信模块 70。该通信模块 70 可以是能够向外部器件(诸如,移动终端、PDA、具有无线 LAN 的便携式终端、支持 USB 或串行通信的个人计算机)传送存储器 50 中存储的三维图像的接口。

[0042] 在下文中,将描述数据存储控制器 40 在存储器 50 中存储三维图像数据的方法。

[0043] 为了有效存储包括立体图像和多个媒体对象的三维图像,根据本发明实施例的存储三维图像数据的方法与基于对象来存储三维图像的方法对应。尤其是,当三维图像包括立体图像和各种媒体对象的组合时(参考图 2),该方法单独地存储三维图像的整体配置性信息、三维图像中包括的每个媒体的对象的类型(即,对象是否为三维图像)、以及用于每个媒体对象的元数据。

[0044] 图 3 图示了根据本发明实施例的三维图像数据的存储格式。参考图 3,由数据存储控制器 40 在存储器 50 中存储的三维图像数据的存储格式包括:ftyp 报头 101、moov 容器 102,和 mdat 容器 103。数据存储控制器 40 在 ftyp 报头 101 中记录文件名、版本信息、文件类型、以及文件的兼容信息。文件类型包括有关三维图像是否包括从立体图像、单视场图像、静态图像、音频图像、及文本媒体之中选择的一个或多个的信息。

[0045] 此外,基于对象编码的三维图像的配置性信息包含在 moov 容器 102 中。而且,用于基于对象编码的三维图像中包括的媒体数据的比特流包含在 mdat 容器 103 中。

[0046] 特别地,moov 容器 102 包括包含用于指示媒体对象间的时间和空间关系的场景描述符信息的字段,其包括:初始对象描述符(IOD)轨道 104、用于场景的二进制格式(BIFS)轨道 105、以及对象描述符(OD)轨道 106。IOD 轨道 104 包括有关 BIFS 轨道和 OD 轨道以及三维图像中包括的所有媒体对象的简档信息。BIFS 轨道 105 包括每个场景中包括的媒体对象的配置性关系的信息。此外,OD 轨道 106 包含有关每个媒体对象的属性的信息。

[0047] 根据如上面描述的本发明实施例,媒体对象的配置性关系的信息包含在 BIFS 轨道 105 中。然而,本发明并不局限于这个实施例,并且包括其中媒体对象的配置性关系的信息可以包含在轨道中的任何实施例。例如,除了 BIFS 轨道 105 之外,媒体对象的配置性关系的信息可包含在轻量级应用场景表达(LASer)轨道中。

[0048] 此外,moov 容器 102 包括:包含三维图像中包括的多个媒体对象的元信息的第一元数据轨道 111、和包含所述多个媒体对象的每个媒体对象中包括的编码流的元信息的第二元数据轨道 112。

[0049] moov 容器 102 包括:包含所述多个媒体对象的每个媒体对象中包括的编码流的配置性信息的流信息轨道 113。在这里,图像的媒体对象可以被不同地定义为立体图像的组合。

[0050] 编码流的配置性信息包括关于配置立体图像和单视场图像的合成图像的配置性信息。

[0051] 其间, OD 轨道 106 包括:包含参考信息的 tref 字段 107。tref 字段 107 包括:不仅包含立体图像、还包含媒体对象和有关与所述媒体对象对应的元数据的对象描述符信息的 ssod 字段 108。

[0052] 特别地, ssod 字段 108 包括:定义配置场景的所述多个媒体对象的媒体对象的数目的信息、以及定义所述多个媒体对象的每个媒体对象是否是立体图像的信息。此外, ssod 字段 108 包括:定义是否存在包含所述多个媒体对象的每个媒体对象的元数据的第一元数据轨道 111 的信息。另外, 当用于每个媒体对象的第一元数据轨道 111 存在时, ssod 字段 108 可进一步包括指示与媒体对象对应的第二元数据轨道 112 的信息。

[0053] 可通过与图 4 所示的程序对应的操作来生成如上面描述的 ssod 字段 108。在图 4 中, 立体 od 入口计数 (stereoscopic\_od\_entry\_count) 指示用于由场景所指示的立体图像的 OD 的数目, 以及立体 od 类型 (stereoscopic\_od\_type) 指示所配置的 OD 是否是立体图像。例如, 当 stereoscopic\_od\_type 被设置为“1”时, 其指示 OD 是立体图像。元数据标志 (metadata-flag) 指示所配置的 OD 是否需要元数据。例如, 当 metadata-flag 被设置为“1”时, 其指示 OD 需要元数据。立体 es 入口计数 (stereoscopic\_es\_entry\_count) 指示所配置的 OD 的数目。元数据轨道标识 (metadata-track-ID) 指示要被应用到所配置的 OD 的元数据的轨道 ID, 以及轨道标识 (track-ID) 指示与 ES 对应的轨道 ID。

[0054] tref 字段 107 进一步包括:包含能够指向第一元容器轨道 111 的信息的 ssom 字段 109。使得访问整体内容的元数据更加便利的 ssom 字段 109 包含指示第一元数据 111 的信息。可以通过与图 5 所示的程序对应的操作来生成如上面描述的 ssom 容器 109。在图 5 中, track-ID 是指用于可应用于整个立体文件的元数据的轨道标识符。

[0055] 此外, mdat 容器 103 包含三维图像中包括的每个媒体对象的编码流。

[0056] 在仅包括立体图像的单一类型内容的情况下, 应以对应的文件格式来描述用于解码组成单一类型内容的立体图像的解码器信息。此外, 在包括立体图像和媒体对象的合成内容的情形下, 一个场景可包括多个媒体对象。因此, 可基于用于多个媒体对象的时间空间配置性信息来存储合成内容。另外, 在基于多个媒体对象的配置性信息来存储合成内容的情况下, 应以文件格式来描述用于根据其每个编码流来解码多个媒体对象的解码器信息。

[0057] 如上所述, 包括不同类型的内容的三维图像需要不同的解码器信息。因此, 为了存储包括不同类型的内容的三维图像, 应为不同类型的内容定义不同的解码器信息。总之, 需要能够支持包括不同类型的內容的所有三维图像的有效的存储方法。

[0058] 在下文中, 将详细描述用于以能够支持包括单一类型内容的三维图像和包括合成内容的三维图像两者的文件格式来存储三维图像的方法。

[0059] 图 6 图示了根据本发明另一实施例的用于存储三维图像数据的格式。参考图 6, 由数据存储控制器 40 在存储器 50 中存储的三维图像数据的存储格式包括: ftyp 报头 201、moov 容器 202、和 mdat 容器 203。

[0060] 根据本发明当前实施例的 ftyp 报头 201 和 mdat 容器 203 分别包含与根据本发明先前实施例的 ftyp 报头 101 和 mdat 容器 103 相同的数据。因此, 这里将省略对 ftyp 报头 201 和 mdat 容器 203 的详细描述。

[0061] 可以基于多个媒体对象来存储包括合成内容的三维图像。当基于多个媒体对象来存储包括合成内容的三维图像时, 数据存储控制器 40 在 moov 容器 202 中装载包含用于

指示多个媒体对象间的时间和空间关系的场景描述符信息的字段,其包括:IOD 轨道 204、BIFS 轨道 205、和 OD 轨道 206。IOD 轨道 204 包括有关 BIFS 轨道和 OD 轨道以及三维图像中包括的多个媒体对象的简档信息。BIFS 轨道 205 包括每个场景中包括的多个媒体对象的配置性关系的信息。此外,OD 轨道 206 包含有关多个媒体对象的每个媒体对象的属性的信息。应当注意到所述多个可以为单一媒体对象,即,至少一个媒体对象。

[0062] 此外,数据存储控制器 40 在 moov 容器 202 中装载对象单元 207 和 208 的多个流轨道,该对象单元 207 和 208 包含多个媒体对象的每个媒体对象中包括的基本流 (ES)。对象单元 207 和 208 的多个流轨道中包括的每个编码流轨道包含用于根据内容的场景描述符的存在或不存在来不同地存储解码器信息的定义。用于不同地存储解码器信息的定义包含在 mpss 容器 209 中。

[0063] 可通过与图 7A、7B、7C、和 7D 所示的示例程序对应的操作来生成 mpss 容器 209。

[0064] 下面的表 1 示出了有关图 7A、7B、7C、和 7D 所示的立体 ES 的详细信息。

[0065] 表 1

| 值    | 内容       |
|------|----------|
| Mpss | 立体可视采样入口 |
| mp4v | 可视采样入口   |
| mp4a | 音频采样入口   |

[0066] 此外,立体组合类型 (StereoScopic\_Composition Type) 包括有关包括一个 ES 的立体数据的图像配置的信息。下面的表 2 示出了信息的指示。

[0067] 表 2

| 值   | 内容                          |
|-----|-----------------------------|
| 000 | 具有根据时间而交叉的帧的立体数据            |
| 001 | 并排 (side by side) 配置的立体数据   |
| 010 | 以由上而下 (top-down) 的方式配置的立体数据 |
| 011 | 具有根据字段而垂直交叉的帧的立体数据          |
| 100 | 具有根据字段而水平交叉的帧的立体数据          |

[0068] 当 StereoScopic\_CompositionType 为“0b001”时,宽 带 或 高 度 (width\_or\_height) 指示图像的宽度;当 StereoScopic\_CompositionType 为“0b010”,其指示图像的高度;当 StereoScopic\_CompositionType 为“0b011”时,其指示每个字段的垂直交叉的垂直线的数目;以及当 StereoScopic\_CompositionType 为“0b100”时,其指示每个字段的水平交叉的水平线的数目。

[0069] 当 StereoScopic\_CompositionType 为“0b001”和“0b010”时,奇数字段计数 (odd\_field\_count) 包括“0”,当 StereoScopic\_CompositionType 为“0b011”和“0b100”时,其包括奇数字段的数目。

[0070] 当 StereoScopic\_CompositionType 为“0b001”和“0b010”时,偶数字段计数 (even\_field\_count) 包括“0”,以及当 StereoScopic\_CompositionType 为“0b011”和“0b100”时,其包括偶数字段的数目。

[0071] 立体 ES 帧同步 (StereoScopic\_ES\_FrameSync) 指示左侧图像 (或奇数字段) 和右侧图像 (或偶数字段) 的帧速率间的差异,且包括用于当帧速率不同时指示同步方法的信息。下面的表 3 示出了通过该信息的指示。

[0074] 表 3

[0075]

| 值  | 内容                                                                                               |
|----|--------------------------------------------------------------------------------------------------|
| 00 | 当立体 ES 计数 (StereoScopic_ES_count) 为 “1” 时, 左侧图像 (或奇数字段) 和右侧图像 (或偶数字段) 的帧速率是相同的。                  |
|    | 当 StereoScopic_ES_count 为 “2” 时, 左侧图像 (或奇数字段) 和右侧图像 (或偶数字段) 的帧速率是相同的或不必要的。                       |
| 01 | 当 StereoScopic_ES_count 为 “1” 时, 由于左侧图像 (或奇数字段) 和右侧图像 (或偶数字段) 的帧速率是不同的, 所以将左侧图像 (或奇数字段) 设置为参考图像。 |
|    | 当 StereoScopic_ES_count 为 “2” 时, 左侧图像 (或奇数字段) 和右侧图像 (或偶数字段) 的帧速率是不同的, 并且基于所声明的 ES 进行同步。          |
| 10 | 当 StereoScopic_ES_count 为 “1” 时, 左侧图像 (或奇数字段) 和右侧图像 (或偶数字段) 的帧速率是不同的, 将右侧图像 (或偶数字段) 设置为参考图像。     |
|    | 当 StereoScopic_ES_count 为 “2” 时, 左侧图像 (或奇数字段) 和右侧图像 (或偶数字段) 的帧速率是不同的, 并且基于匹配的 ES 进行同步。           |

[0072] 立体图像信息差异 (StereoScopic\_ImageInformationDifference) 包括用于指示左侧图像 (或奇数字段) 和右侧图像 (或偶数字段) 的图像信息间的图像信息的差异的信息。下面的表 4 示出了通过该信息的指示。

[0073] 表 4

[0074]

| 值  | 内容                                                                      |
|----|-------------------------------------------------------------------------|
| 00 | 当 StereoScopic_ES_count 为 “1” 时, 左侧图像 (或奇数字段) 和右侧图像 (或偶数字段) 间的图像信息没有差异。 |
|    | 当 StereoScopic_ES_count 为 “2” 时, 左侧图像 (或奇数字段) 和右侧                       |

[0075]

|    |                                                                                       |
|----|---------------------------------------------------------------------------------------|
|    | 图像（或偶数字段）间的图像信息不存在或不需要差异。                                                             |
| 01 | 当 StereoScopic_ES_count 为“1”时，左侧图像（或奇数字段）和右侧图像（或偶数字段）间的图像信息存在差异，且基于左侧图像（奇数字段）来获得该差异。  |
|    | 当 StereoScopic_ES_count 为“2”时，左侧图像（或奇数字段）和右侧图像（或偶数字段）间的图像信息存在差异，且基于所声明的 ES 来获得该差异。    |
| 10 | 当 StereoScopic_ES_count 为“1”时，左侧图像（或奇数字段）和右侧图像（或偶数字段）间的图像信息存在差异，且基于右侧图像（或偶数字段）来获得该差异。 |
|    | 当 StereoScopic_ES_count 为“2”时，左侧图像（或奇数字段）和右侧图像（或偶数字段）间的图像信息存在差异，且基于匹配的 ES 来获得该差异。     |

[0076] Y或R差异 (Y\_or R difference) 指示图像信息 Y 差值或 R 差值, Cb 或 G 差异 (Cb\_or\_G\_difference) 指示图像信息 Cb 差值或 G 差值, 且 Cr 或 B 差异 (Cr\_or\_B difference) 指示图像信息 Cr 差值或 B 差值。

[0077] 此外, 立体 ES 类型 (StereoScopic\_ES\_Type) 指示有关 ES 的立体类型的信息。下面的表 5 示出了那个信息。

[0078] 表 5

[0079]

| 值  | 内容   |
|----|------|
| 00 | 左侧图像 |
| 01 | 右侧图像 |
| 10 | 参考图像 |
| 11 | 附加数据 |

[0080] 流类型 (StreamType) 指示 ES 流的类型, 且与 14496-1 中 ES 描述符 (ES\_Descriptor) 的解码器配置描述符 (DecoderConfigDescriptor) 的 StreamType 执行相同的功能 (参见 ISO 基本媒体文件格式的 14496-18.6.6 的“StreamTypeValues (流类型值)”表)。

[0081] 对象类型指示 (ObjectTypeIndication) 指示 ES 流的必需的解码器, 其与 14496-1 中 ES\_Descriptor 的 DecoderConfigDescriptor 的 ObjectTypeIndication 执行相同的功能 (参见 ISO 基本媒体文件格式的 14496-18.6.6 的“ObjectTypeIndication Values (对象类型指示值)”表)。

[0082] 解码器特定信息长度 (DecoderSpecificInfoLength) 指示解码器特定信息流 (DecoderSpecificInfoStream) 的长度, 且与 14496-1 中 ES\_Descriptor 的 DecoderSpecificInfoLength 执行相同的功能。

[0083] 其间, 图 8 是图示了根据本发明另一实施例的在包括场景描述符信息的情况下的

数据存储格式的结构,而图 9 图示了根据本发明另一实施例的在不包括场景描述符的情况下的数据存储格式的结构。

[0084] 如图 8 所示,当具有合成内容的三维图像包括场景描述符信息时,该三维图像不可避免地包括 IOD 信息。因此,通过确定它是否包括包含 IOD 信息的 IOD 字段,可能确定该三维图像是否包括场景描述符信息。因此,如果 moov 容器包含 IOD 信息,则有关三维图像的解码器信息存储在 MPEG-4 系统中定义的 ES 描述符 (ESD) 中。

[0085] 相反地,如图 9 所示,当具有单一类型内容的三维图像不包括 IOD 信息时,有关三维图像的解码器信息存储在 mpss 字段中。存储在 mpss 字段中的解码器信息包括:流类型、流的媒体对象的类型(例如,MPEG-4 系统中定义的对象类型 (Object type) 指示)、简档信息、以及解码器还原媒体对象所需的详细信息(例如,图像的宽度)。

[0086] 通过与图 7A、7B、7C、和 7D 所示的程序对应的操作,尤其是通过图 7B 的操作 300 和图 7D 的操作 350,数据存储控制器 40 可以考虑场景描述符号信息的存在或不存在而有效地存储有关三维图像的解码器信息。

[0087] 另外,数据存储控制器 40 装载包含有关 moov 容器 212 中的对应的元数据的信息的元信息字段 210。元信息字段 210 包括 iloc 字段 211 和 iinf 字段 212。iloc 字段 211 包含编码流的标识符(例如,Item ID)、和有关该编码流的比特流的长度的信息(例如,偏移和长度),而 iinf 字段 212 包含每个编码流的名称、内容类型、和编码信息。

[0088] 尽管已经参考本发明的特定示范实施例示出和描述了本发明,但是本领域技术人员将理解,可以在其中进行形式和细节上的各种改变,而不脱离通过所附权利要求所限定的本发明的精神和范围。

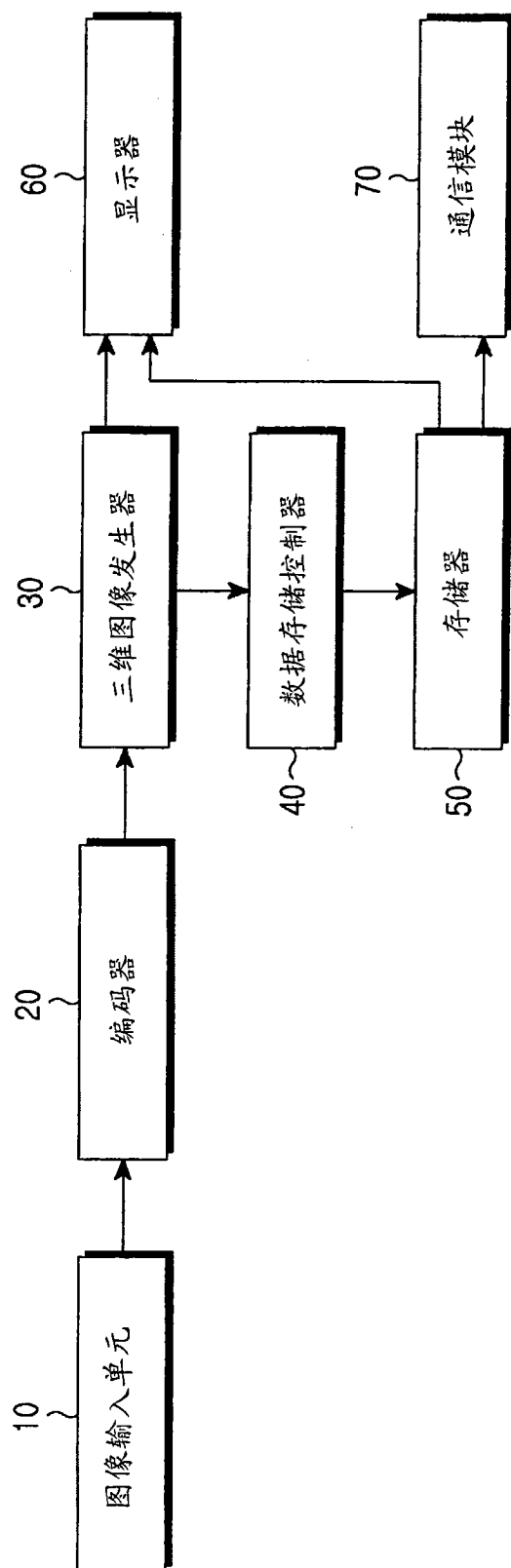
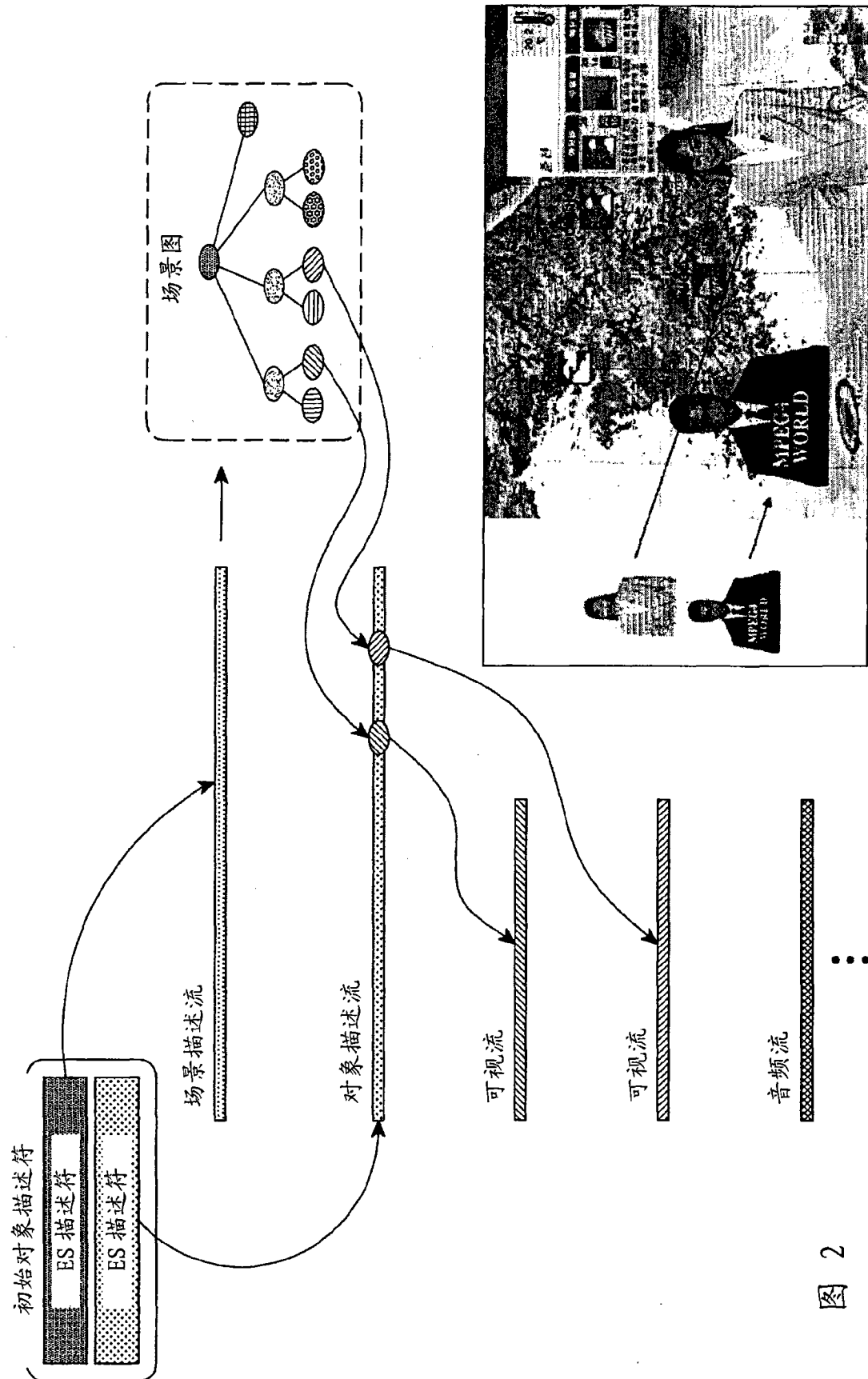


图 1



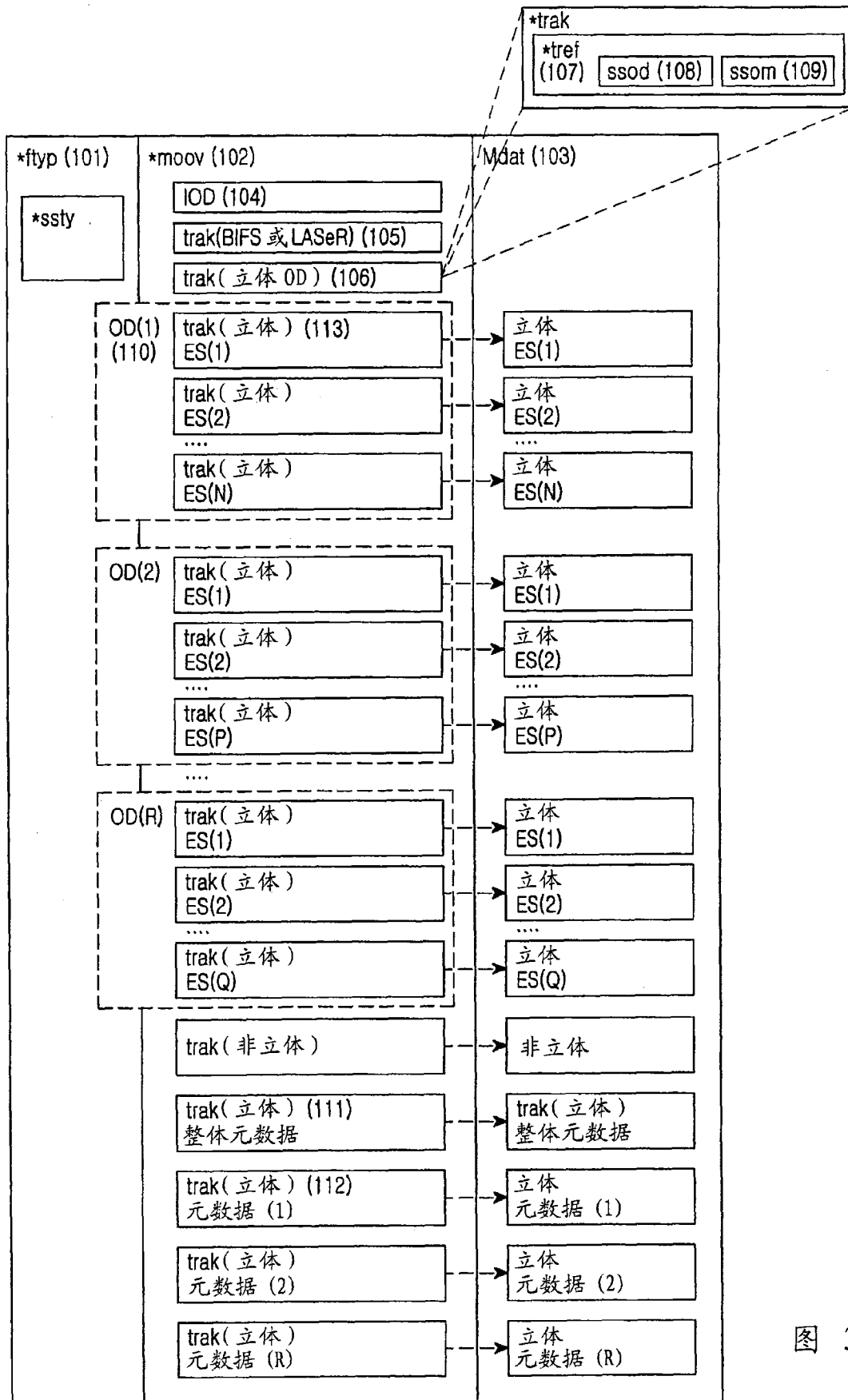


图 3

### ssod (立体对象描述符类型)

逻辑单元类型: “ssod”

容器: 轨道参考逻辑单元 (“tref”)

强制性: 否

数量: 零个或一个

#### 语法

```
aligned (8) class StereoScopicObjectDescriptorTypeBox extend Box ('ssod', 0, 0) {
  int l, j;
  unsigned int (32) stereoscopic_od_entry_count;
  for( i=0; i<stereoscopic_od_entry_count; i++)
  {
    unsigned int (1) stereoscopic_od_type;
    unsigned int (1) metadata-flag;
    unsigned int (6) reserved;

    unsigned int (32) stereoscopic_es_entry_count;
    if ( metadata-flag == 1 ) {
      unsigned int (32) metadata-track-ID;
    }

    for ( j=0; j<stereoscopic_es_entry_count; j++)
    {
      unsigned int (32) track-ID[j];
    }
  }
}
```

图 4

### ssom (立体整体元数据类型)

逻辑单元类型: “ssom”

容器: 轨道参考逻辑单元 (“tref”)

强制性: 否

数量: 零个或一个

#### 语法

```
aligned (8) class StereoScopicOverallMetadataTypeBox extend Box ('ssom', 0, 0) {
  unsigned int (32) track-ID;
}
```

图 5

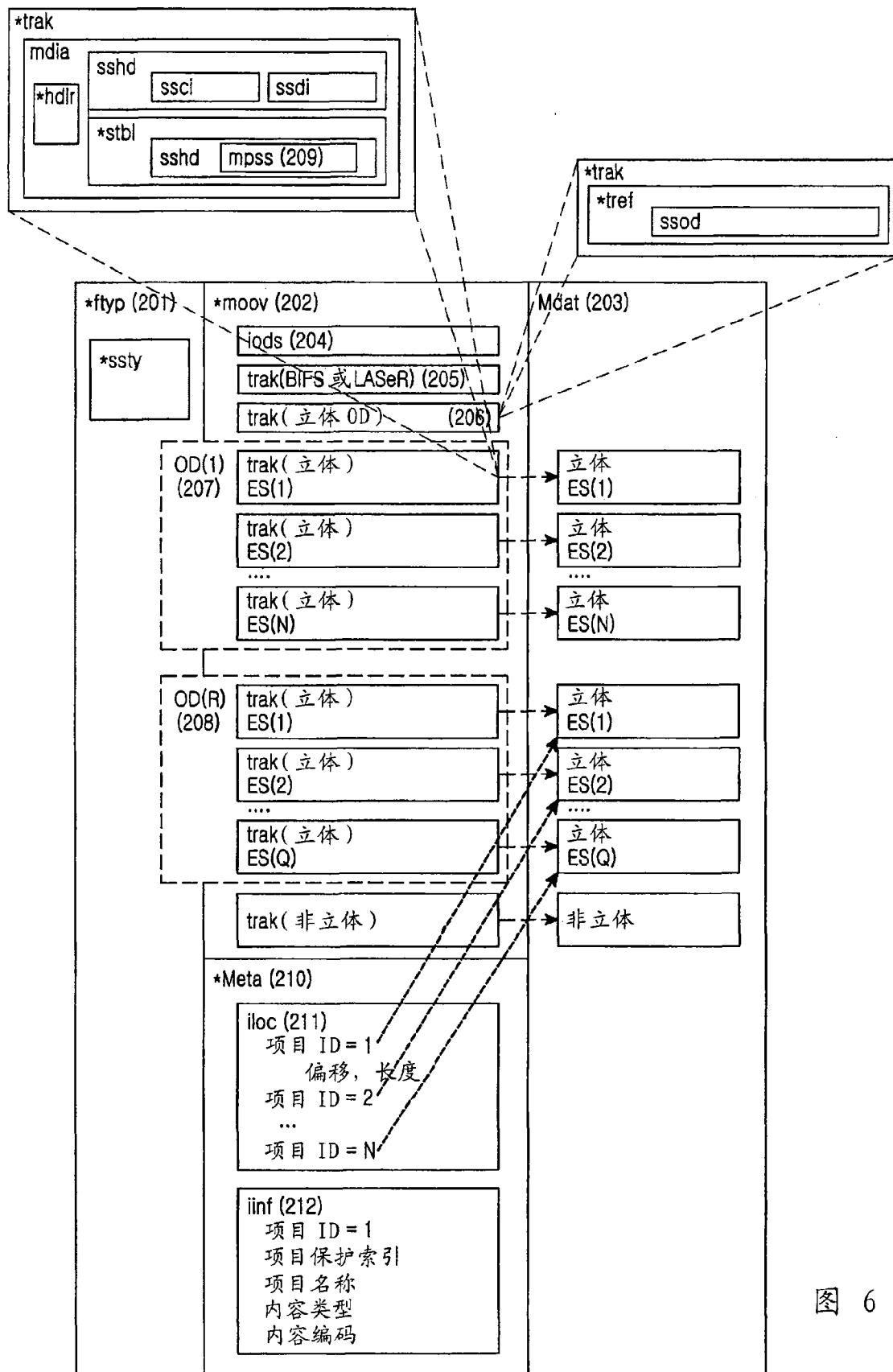


图 6

mpss ( 立体可视采样入口 )

逻辑单元类型: “mpss”、“mp4v”、“mp4a”、“mp4s”

容器: 采样表逻辑单元 ( “stbl” )

强制性: 是

数量: 正好为一个

语法

```
aligned (8) class ESDBox extends Box ('esds', version = 0, 0) {  
    ES_Descriptor ES; // 参考 14496-1  
}  
aligned (8) abstract class SampleEntry (unsigned int (32) format) extends Box (format) {  
    // 参考 14496-12 ISO 基本媒体文件格式  
}  
class HintSampleEntry () extends SampleEntry (protocol) {  
    // 参考 14496-12 ISO 基本媒体文件格式  
}  
class AudioSampleEntry (codingname) extends SampleEntry (codingname) {  
    // 参考 14496-12 ISO 基本媒体文件格式  
}  
class VisualSampleEntry (codingname) extends SampleEntry (codingname) {  
    // 参考 14496-12 ISO 基本媒体文件格式  
}  
class MpegSampleEntry () extends SampleEntry ('mp4s') {  
    ESDBox ES; // 参考 14496-1  
}
```

图 7A

```

class MpegVisualSampleEntry () extends VisualSampleEntry ('mp4v') {
    ESDBox ES; // 参考 14496-1
}

```

300

```

class Mpeg Audio Sample Entry () extends Sample Entry ('mp4a'){
    if ( iod_track ) {
        ESDBox ES;
    } else {
        unsigned int (8)  StreamType;
        unsigned int (8)  ObjectTypeInfo;
        unsigned int (8)  DecoderSpecificLength;
        for ( int i=0; i<DecoderSpecificLength; i++ )
            unsigned int (8)  DecoderSpecificInfoStream[i];
    }
}
class StereoScopicVisualSampleEntry(unsigned int(4) StereoScopic_ES_count) extends
VisualSampleEntry ('mpss') {

```

```

    if ( StereoScopic_ES_count == 1 ) {
        unsigned int (3)  StereoScopic_CompositionType;
        unsigned int (1)  reserved;
        unsigned int (2)  StereoScopic_ES_FrameSync;
        unsigned int (2)  StereoScopic_ImageInformationDifference;
        if ( StereoScopic_CompositionType != 0 ) {
            int (16)  width_or_height;
            int (16)  odd_field_count;
            int (16)  even_field_count;

```

图 7B

```
    if ( StereoScopic_ImageInformationDifference != 0 ) {  
        int (16)    Y_or_R_difference;  
        int (16)    Cb_or_G_difference;  
        int (16)    Cr_or_B_difference;  
    }  
} else if ( StereoScopic_ES_count == 2 ) {  
    unsigned int (2)    StereoScopic_ES_Type;  
    unsigned int (2)    reserved;  
    unsigned int (2)    StereoScopic_ES_FrameSync;  
    unsigned int (2)    StereoScopic_ImageInformationDifference;  
  
    if ( StereoScopic_ImageInformationDifference == '0b10' ) {  
        int (16)    Y_or_R_difference;  
        int (16)    Cb_or_G_difference;  
        int (16)    Cr_or_B_difference;  
    }  
} else {  
    unsigned int (2)    StereoScopic_ES_Type;  
    unsigned int (6)    reserved;  
}
```

图 7C

```

350
}
if ( iod_track ) {
    ESDBox ES;
} else {
    unsigned int (8)    StreamType;
    unsigned int (8)    ObjectTypeInfoIndication;
    unsigned int (8)    DecoderSpecificInfoLength;
    for ( int i=0; i<DecoderSpecInfoLength; i++ )
        unsigned int (8)    DecoderSpecificInfoStream[i];
    }
}

class SampleDescriptionBox (unsigned int (32) handle-type) extend Box ('std', 0, 0) {
    int i;
    unsigned int (32)    entry-count;
    for ( i=0; i<entry-count; i++ ) {
        switch (handler-type) {
            case 'soun' : // 音频流
                MpegAudioSampleEntry();
                break;
            case 'vide' : // 可视流
                MpegVisualSampleEntry();
                break;
            case 'ssvi' : // 立体可视流
                StereroScopicVisualSampleEntry();
                break;
            case 'hint' : // 提示轨道
                HintSampleEntry();
                break;
            default :
                MpegSampleEntry();
                break;
        }
    }
}

```

图 7D

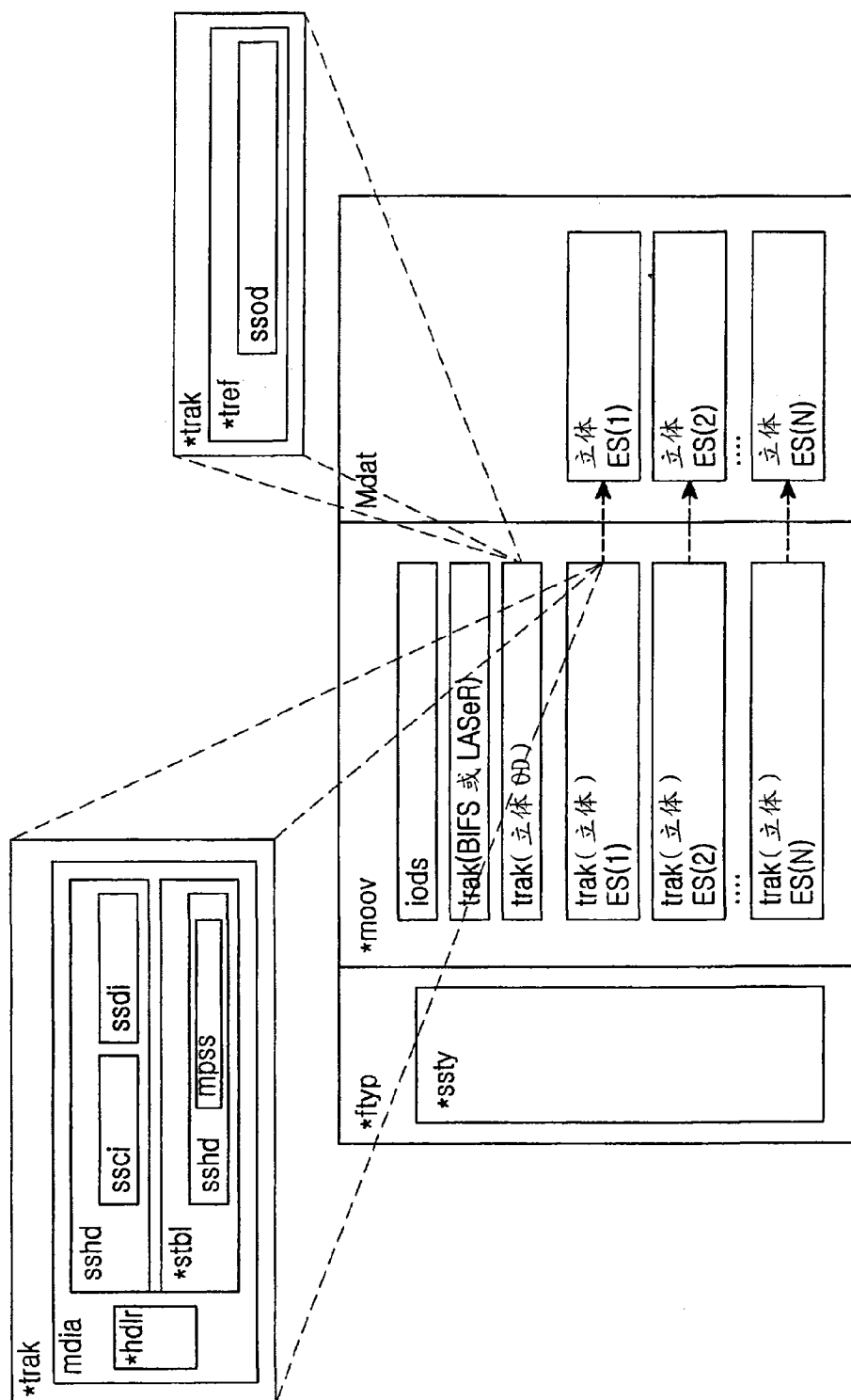


图 8

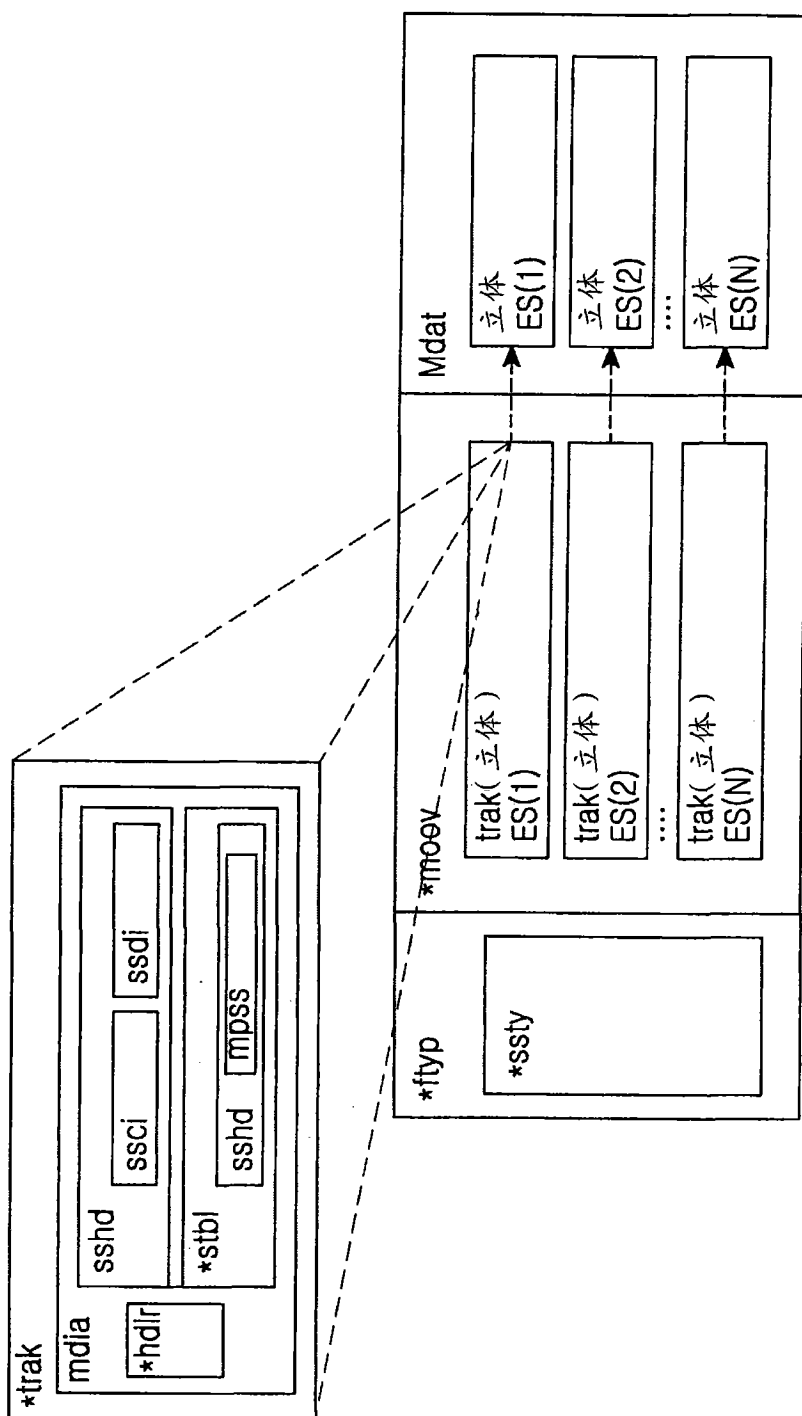


图 9